The Journal of National Chiao Tung University
Vol. 6, April 1979
pp. 23-29

—— 短文 *Short Notes*

中文符傳系統之設計與建立
Design and Implementation of A Chinese FORTRAN System.

蔡中川 · 蔡汪銓 Jong-Chuang Tsay and Wang-Chuan Tsai

Institute of Computer Science

(Received May 3, 1978)

Abstract — The design and implementation of a Chinese FORTRAN System are presented. This Chinese FORTRAN System preserves the full capabilities of a FORTRAN language with the added features of representing variable name and data in Chinese and performing Chinese I/0 operations. Besides, a Chinese tabulation language, CHITAL [1] has been incorporated into the Chinese FORTRAN System to enhance the capability of controlling output format of Chinese data. A sample program and its running result are given in the end for reference.

Due to the success of the development of Chinese I/O system, the idea of implementing a Chinese high-level language has been widely discussed [2] [3] [4]. In this paper, a new Chinese high-level language, Chinese FORTRAN, is presented. It is developed on PDP 11/40 of NCTU. Basically, the syntax structure of our Chinese FORTRAN language is based on the FORTRAN language of PDP 11/40 DOS/BATCH system [5]. However, since several features are added specifically for the Chinese data processing, the original syntax structure should be modified slightly. These added features are:

(1) Since Chinese FORTRAN program is created by using Chinese Text Editor [6] [7] via Chinese keyboard, Chinese characters can appear in our Chinese FORTRAN program.

(2) A variable can use up to three Chinese characters as its name. For example: 姓名 (16)，總收入，加B班，I 合計。

(3) A special Chinese DATA statement (CDATA) is designed to specify Chinese data and/or CHITAL commands [1]. It must precede DATA statement and any executable statement but follow any other specification statements. The followings are examples of CDATA statement:

CDATA FM1/4X, " 國立交通大學 ", E/
CDATA FM2/8(ϕJ.3J, 576S, ϕS)/

where FM1, FM2 are the variables of the CDATA statement or in short, the CDATA variables. The character strings between "/ " are called CDATA data string.

(4) A variable can be used to store Chinese character string. The type and dimension of such a variable are BYTE and 4*(n+1) where n is the maximum number of Chinese characters the variable can store.

For example: Let n=3 for the variable 姓名 then we should use the following type declaration

BYTE 姓名 (16)

(5) CDATA variable may appear in the I/O list of WRITE statement. Under this circumstance, the corresponding FORMAT statement should follow the WRITE statement immediately and the field descriptor in the format specification for the variable should be written as ??A4.

For example:

```
      CDATA FM1/......../
      CDATA FM3/......../
        .
        .
        .
      WRITE (3,15) FM1, 姓名, FM3, VAR
   15 FORMAT(1X, ??A4, 16A1, ??A4, F5.1)
        .
        .
        .
```

(6) The Chinese character string in the Chinese input data file must be enclosed by <and >. For example, the Chinese character string for the variable 姓名 (16) should appear in the input data file as

····<蔡中川>····

With the Chinese FORTRAN language as described above we can now go on to discuss the system which is designed to process the language.

The Chinese FORTRAN System is implemented on the PDP 11/40 of NCTU. In respect of software, the Chinese FORTRAN System has a block diagram of Fig. 1.

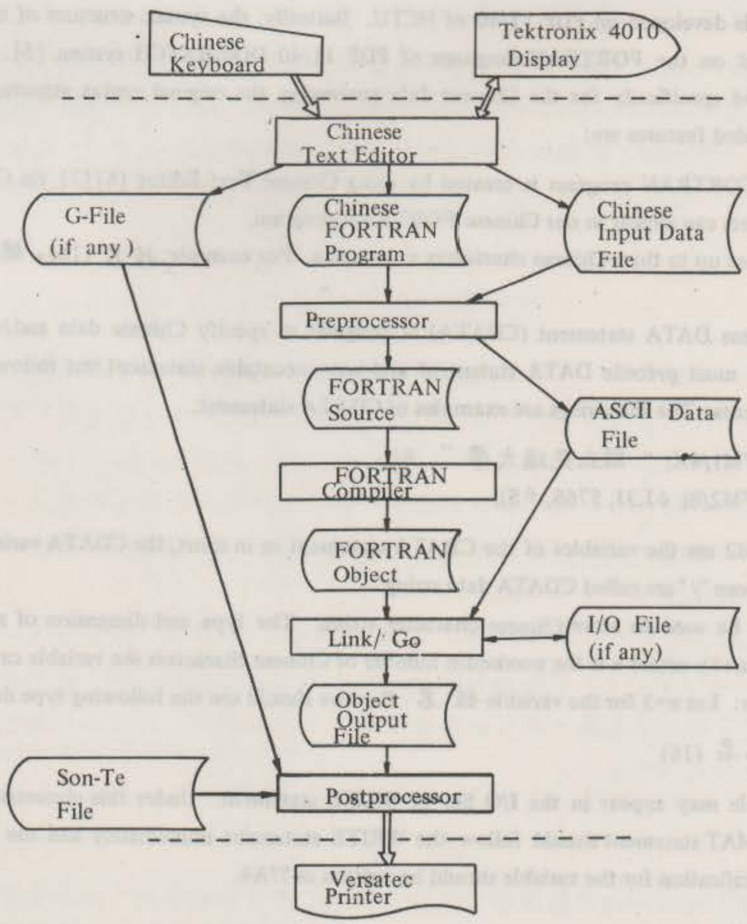

Fig. 1 System Block Diagram

By using a Chinese text editor [6] [7], we can create the Chinese FORTRAN program file or Chinese input data file via Chinese keyboard and graphic display. Since these files are formatted binary files, they should be preprocessed into a compilable FORTRAN source and readable ASCII data file respectively. The Preprocessor read in the record of Chinese FORTRAN program one-by-one and process it according to the following steps:

(1) Change the formatted binary record into an ASCII record.

. For a real (integer) variable with Chinese name, it is replaced by the variable XXnnn(IXXnnn), where nnn is a 3-digit number generated by the preprocessor.

. For a Chinese character string which is enclosed by double apostrophe and presented in the CDATA statement, it is handled by changing each Chinese character into four ASCII codes. For example, the following statement.

CDATA FM/........ "合計".......∕

is changed into

CDATA FM/...."ED12 2C32"...../

where $21DE_{16}$, $23C2_{16}$ are internal codes for the Chinese characters 合計 .

(2) If it is a CDATA statement, we need to do the following:

. Enclose the CDATA data string by preceding its first character with the character < and following its last character with > . The whole data string is then partitioned into elements. Each element cosists of four characters. The number of elements in the data string is then used as the dimension of the CDATA variable.

. This CDATA variable is then declared REAL.

For example

CDATA FM1/3X, "ED12 2C32 A3FF"/

is translated into

REAL FM1(5)

DATA FM1/ '<3X,' , '"ED1', '22C3', '2A3F', 'F"> '/

(3) If it is a FORTRAN statement and its statement number equals to the format specifier of previous WRITE statement, the unknown repeat count?? in the format specification of FORMAT statement will be replaced by the dimension of the corresponding CDATA variable of WRITE statement.

For a Chinese input data file, the Preprocessor change each formatted binary record into an ASCII record. Chinese character string, which is enclosed by <and> is handled by changing each Chinese character into four ASCII codes and enclosing the full string in double apostrophe. For example, the Chinese input data file

< 蔡中川 > 1 φφ.5

is translated into an ASCII file of

<"C326 FB φφ 96 φφ " >1 φφ.5

where $623C_{16}$, $\phi\phi BF_{16}$, $\phi\phi 69_{16}$ are internal codes of the Chinese characters

Refer to Fig. 1 again, the Chinese FORTRAN program and Chinese input data file is translated by the Preprocessor into FORTRAN source and ASCII data file respectively. The FORTRAN source is compiled by the FORTRAN compiler to generate a FORTRAN object file. The object file is linked and executed to generate an object output file which is an ASCII file. It is postprocessed to list the final result on Versatec printer. The Postprocessor consists of two parts: output file processor and CHITAL processor [1]. Output file processor accepts the records from object output file and then generates two files for the input of CHITAL processor. One is called A-file and the other is an ASCII CHITAL command string [1]. These files are generated in the following way:

(1) Those characters enclosed by double apostrophe are packed into Chinese internal code and stored in the A-file sequentially. An nA CHITAL command is then added to the CHITAL command string, where n is the number of Chinese characters.

(2) The calculated data part of the object output file are written directly to the A-file sequentially. An nA CHITAL command is then added to the CHITAL command string, where n is the number of characters in the calculated data.

For example, the object output file

<2X, "C326FBφφ 96φφ " >1φφ.5

generates

CHITAL command string: 2X, 3A, 5A, E

A-file: 蔡中川 1φφ.5

The CHITAL command string and A-file can be passed to the CHITAL processor of the Postprocessor for listing the final result. Besides A-file, G-file (see Fig. 1) is an optional input file for the CHITAL processor. It has a file format same as A-file and can be read by including an nG command in the CHITAL command string. See Reference [1] for the details of the CHITAL processor. After having described the Chinese FORTRAN system, we can now present a sample program and its result to illustrate the usage of our system. This is a students' score reporting program. Ten students' names and their scores of five courses are read from a Chinese input data file. Then the total score, total credit, and the average score of each student are calculated. The tabulation control is defined in the CDATA statements. The title of the table is read from the G-file.

Figure are arranged as follows:

Figure	Description
2	Chinese FORTRAN program
3	Chinese input data file
4	G-file
5	Listing output

```
      BYTE 姓名(10.16)
      IMPLICIT INTEGER (X)
      DIMENSION 學分(5).成績(10.5).總分(10).總學分(10)
      DIMENSION AVE(10).名次(10)
      CDATA TITL/6P2.R(2X.18G.E).8L.6P1.*/
      CDATA TL1/72D.6T.-2B.2D.6(3T.-2B.2D).4T
     *.-2B.2D.2T.-2B.2D.4T.-2B.2D.2T.-2B.2D.
     *6T.18B.72D.*/
      CDATA TL2/2(2X.G).5(X.G.X).4X.18B.3(G.2X
     *).G.3X.G.E.3X.G.X.G.22X.18B.G.E.6X.*/
      CDATA TL3/5(X.G.X).4X.18B.3(G.2X).G.3X.
     *G.E.2X.2G.2X.*/
      CDATA TL4/"(4)(4)(3)(3)(3)".E.8L/
     *VLNE/72D.8T.-2B.2D.5(3T.-2B.2D).3(18D.
     *T).-2B.2D.2(4T.-2B.2D.2T.-2B.2D).6T.18
     *B.72D.*/HLNE/8L/
      CDATA ENDL/6P2.R(8X."三年15班".E).6E/
      DATA 學分/4.4.3.3.3/
      CALL ASSIGN(3.'LIST.DAT'.8)
      CALL ASSIGN(2.'NUM DAT'.7)
      WRITE(3.100)TITL.TL1.TL2.TL3.TL4
  100 FORMAT(??A4/??A4/??A4/??A4/??A4)
      DO 104 I=1.10
      READ(2.101)(姓名(I.T).T=1.16)
  101 FORMAT(16A1)
```

```
      READ(2,102)(成績(I,T),T=1,5)
  102 FORMAT(5I3)
      總學分(I)=0
      總分(I)=0
      DO 103 T=1,5
      IF(成績(I,T).LT.60)GOTO 103
      總學分(I)=總學分(I)+學分(T)
  103 總分(I)=總分(I)+成績(I,T)*學分(T)
  104 AVE(I)=總分(I)/17.
      DO 111 I=1,10
      名次(I)=0
      DO 110 T=1,10
      IF(AVE(I).GE.AVE(T))GOTO 110
      名次(I)=名次(I)+1
  110 CONTINUE
  111 名次(I)=名次(I)+1
      DO 120 I=1,10
      WRITE(3,112)VLNE,(姓名(I,T),T=1,16),
     *(成績(I,T),T=1,5),總分(I),總學分(I),AVE(I),
     *名次(I)
  112 FORMAT(??A4/2X,16A1,1X,5I3,3X,I4,I2,
     *F4.1,I2)
      WRITE(3,113)HLNE
  113 FORMAT(??A4)
  120 CONTINUE
      WRITE(3,121)ENDL
  121 FORMAT(??A4)
      END FILE 2
      END FILE 3
      STOP
      END
```

Fig. 2

```
〈林竹林〉
  90 70 85 65 69
〈南宮申〉
  78 87 88 77 88
〈白珀琥〉
  84 85 86 87 88
〈楊萌兆〉
  90 69 85 66 77
〈包  重〉
  55 59 73 20 60
〈古存AA〉
  91 76 75 48 64
〈林武牙〉
  74 71 92 58 76
〈黃  河〉
  80 81 82 83 84
〈崔九肅〉
  68 58 78 65 80
〈來磐川〉
  59 76 95 86 67
```

Fig. 3

```
ØØØØ   台北市立××中學第_次月考成績通知單
ØØØ1   成科國數英物化總總平名註
ØØØ2   績目學文學文理學分分均次脚姓名
```

Fig. 4

台北市立ＸＸ中學第一次月考成績通知單

成績 科目 姓名	國文 (4)	數學 (4)	英文 (3)	物理 (3)	化學 (3)		總分	總學分	平均	名次	註脚
林竹林	90	70	85	65	69		1297	17	76・3	5	
南宮申	78	87	88	77	88		1419	17	83・5	2	
白珀琥	84	85	86	87	88		1459	17	85・8	1	
楊萌兆	90	69	85	66	77		1320	17	77・6	4	
包　重	55	59	73	20	60		915	6	53・8	10	
古存此	91	76	75	48	64		1229	14	72・3	8	
林武牙	74	71	92	58	76		1258	14	74・0	7	
黃　河	80	81	82	83	84		1391	17	81・8	3	
崔九肅	68	58	78	65	80		1173	13	69・0	9	
來譬川	59	76	95	86	67		1284	13	75・5	6	

三年15班

Fig. 5

The implementation of a Chinese FORTRAN System has not been done before. In this paper, we present an easy method to implement such a system. As can be seen from the sample program the implemented system keeps all the capabilities of a FORTRAN language with additional power of doing Chinese I/O operations and tabulating Chinese characters.

Acknowledgment

The authors would like to thank National Science Council, ROC, for the financial support. The authors wish to thank those heros who have contributed to the development of Chinese I/O system on PDP11/40 of NCTU.

References

1. J. C. Tsay, "A new language for tabubtion-CHITAL," The Jol. of NCTU, 1, 41-56, (1976)
2. 趙崇耀，蒲浩哲，"中文基礎程式語言之研究", Proc. of the Nat'l Computer Symposium of the ROC, Taiwan University, 8-27–8-28, (1976)
3. Yaohan Chu, "A Chinese high-level language research computer system", Proc. of Int'l Computer Symposium, Vol. I, 199-203, (1975)
4. C. C. Lee, W. H. Cheng, Y. C. Chen, and H. S. Liang, "The design and implementation of Chinese Report Program Generator",

Vol. I, 179-185, (1975)

5. "The DOS/BATCH Handbook", Part 7, The DOS/BATCH FORTRAN Compiler and Object Time System, DEC, 1973.

6. 蔡中川，"中文編輯系統之設計及其實例說明"，計算機園地，創刊號，1－9頁 66年2月。

7. 陳永健，蔡中川，"計算機輔助之中文印刷系統"，科學發展月刊，第五卷第七期 558－569頁，66年7月。